Tilburg University

Inventory Control

Strijbosch, L.W.G.; Moors, J.J.A.

Publication date:
1998

Citation for published version (APA):
Strijbosch, L. W. G., & Moors, J. J. A. (1998). *Inventory Control: The Impact of Unknown Demand Distribution*. (FEW Research Memorandum; Vol. 770). Econometrics.

INVENTORY CONTROL: THE IMPACT OF UNKNOWN DEMAND DISTRIBUTION

L.W.G. Strijbosch[1]

J.J.A. Moors[1]

Abstract

A (R, S) inventory control policy is considered with fixed review time R. The order-up-to-level S has to satisfy one of the following two service criteria:

- the P_1-criterion, requiring that the stock-out probability is at most $1-P_1$,
- the P_2-criterion, requiring that the fraction of ordered quantities that can not be delivered from stock is at most $1-P_2$.

As far as possible, these well-known criteria will be treated simultaneously and uniformly. For simplicity, only zero leadtimes are considered throughout the paper.

For stationary, normally distributed demand with known parameters μ and σ, the safety factors c_i and order-up-to-levels S_i for the P_i-criterion ($i=1,2$) can be found in most textbooks on inventory control. In case of unknown parameters, the standard literature advocates the same formulae, but with μ and σ replaced by estimates. It is shown here that the required service level P_i is not met, when this standard procedure is followed; in typical cases, the failure probabilities may exceed the prescribed values $1-P_i$ with 10-30%. Simple explanations of this phenomenon are given.

An obvious remedy is to enlarge the safety factors. For known coefficient of variation ν, the safety factors leading *exactly* to the prescribed service levels were found by simulation. It appears that for normally distributed demand, c_i should be increased up to 7%; for stationary gamma distributed demand, c_1 should be increased up to 80% and c_2 even up to 175%!

Since ν is unknown in practice, these exact safety factors are not applicable automati-

[1]Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands; email-address first (corresponding) author: L.W.G.Strijbosch@Kub.nl

cally. Hence time-varying safety factors were developed, based upon exponential smoothing. Their behaviour is studied in detail, first for normal demand: in stationary situations, they lead to service levels only slightly below the desired values.

For gamma distributed demand however, even these newly developed safety factors fall short: in typical cases the stock-out probability can be twice the intended value, the fraction of orders that can not be delivered from stock even may be tripled. Hence, still another type of safety factor was developed, depending on past performance. It achieves service levels closer to the prescribed values - although its practical use may be limited.

Key words:

fill rate performance, gamma distribution, normal distribution, P_1-service level, P_2-service level, (R,S)-control policy, safety factor, service criterion, simple exponential smoothing, simulation, standard rule, stock-out probability.

1. Introduction

Inventory control is one of the most intriguing problems in logistics - both from a theoretical and from a practical point of view. An abundance of papers is devoted to this large research area. Essentially, inventory control considers two questions:

- how can inventory be kept at a sufficiently high level?,
- what does ‘sufficiently’ exactly mean here?

To answer the first question, many inventory control systems have been suggested in the literature. Here, we stick to one of them: throughout the paper, the familiar (R,S)-control system will be considered. In this system, the inventory position is measured at certain moments, and replenished to level S. The time interval between two measurements is the review period R. We will assume that R is a given constant; in addition lead-time will be assumed to be zero. The (R,S) control policy is widely applicable. One reason for its usage is the fact that the ordering and the delivering organisations can make clear appointments on the order and delivery dates. An example of a situation where an (R,S) policy with zero lead times is appropriate is the inventory control of medicines in a hospital department, where at the end of a (possible drug-dependent) number of weeks, the inventory of drug *i* is replenished up to a level S_i within (for example) one hour. Of course, the zero lead-time assumption severely limits the generality of our results; we will show, however, that even this restricted problem has many interesting features.

With respect to the second question, the values for the decision parameters (S in the (R,S)-policy) are commonly determined with help of either a cost or a service criterion. In this paper we concentrate on two different service criteria:

- the fraction of *review periods* in which total demand can be delivered from stock should be equal to a given (minimum) level P_1,
- the fraction of *total demand* that can be delivered from stock should equal a given (minimum) level P_2.

These criteria are called the P_1- and P_2-service criterion, respectively; the latter is also known as the fill rate performance. Our approach aims at a general treatment for both criteria.

To tackle the problem of inventory control analytically, assumptions are needed about the behaviour of demand during a review period (review plus lead-time, when lead-time is larger than zero). First of all, demand is assumed to be stationary, that is, the parameters of the distribution do not change over time. Let X_t denote demand during one review period; N denotes a normal and Γ a gamma distribution. Three different assumptions about the distribution of X_t will be used in the sequel:

- $X_t \sim N(\mu, \sigma^2)$ with known μ and σ,
- $X_t \sim N(\mu, \sigma^2)$ with unknown μ and/or σ,
- $X_t \sim \Gamma(\lambda, \rho)$ with unknown λ and ρ.

In the last two cases, estimation procedures have to be selected. Since stationary time series are considered, the best estimators for mean and variance are of course based equally on all past observations. In practice, however, changes in demand distribution will occur. To allow for these changes, it is usual to estimate μ and σ by simple exponential smoothing (SES). We shall follow this practice in our paper. In other words: methods developed for non-stationary situations will be evaluated in stationary cases only.

Most papers on inventory control with unknown demand distribution focus on the relative performance of several forecasting methods with respect to attained service levels or inventory investments. The most important instrument of analysis is Monte Carlo experimentation. Jacobs and Wagner (1989) investigate the impact on total system cost of using the sample mean and standard deviation as compared to robust parameter estimates, like exponentially smoothed average and mean absolute deviation (MAD). An important outcome of their research is that, in general, the scaling factor of the MAD should be larger than the commonly used 1.25. Gardner (1990) compares the influence of several forecasting techniques on the relation between customer service and inventory investment. By means of simulation, Watson (1987) showed that for lumpy (stationary stuttering Poisson) demand patterns, demand-forecast fluctuation (using Simple Moving Average) can cause either positive or negative shifts in the customer service level achieved. Karmarkar (1994) presents an alternative method for estimating the service level, in stead of the conventional approach based on the normal model. His method is especially suited to meet situations where the type of the demand distribution is unknown. Eppen and Martin (1988) very clearly describe some consequences of incorrectly assuming normality of the distribution of forecast errors over lead-time. Strijbosch et al. (1998) investigate by simulation the performances of a simple (more or less standard) control policy and a more advanced one, based on the compound Bernoulli model. For a (s, Q)-control policy with a cost criterion, Silver and Rahnama (1987) showed that underestimating the safety factor k leads to a higher cost penalty than overestimating k; they conclude therefore that deliberately biasing k upwards will reduce costs.

This result of Silver and Rahnama (1987) deserves more attention. We analyse this issue here in the setting of an (R, S)-policy with stationary demand with unknown parameters, zero lead-time, and the two service criteria P_1 and P_2. As an introduction, the very simple case of a fully known normal demand distribution is considered in Section 2. For the two service criteria mentioned before, standard theory offers explicit formulae for S; e.g., see Silver et al. (1998); they can be combined into one expression, having as general feature that S should exceed μ by a certain multiple of σ. Standard theory advocates to apply the same multiplication factors in the case that μ and σ are replaced by estimates (the *standard rule*). However, it is shown theoretically in Section 3, that the desired service levels are not attained by applying this standard rule: once again, it appears that the safety factors should be increased.

Section 4 investigates by Monte Carlo experiments the service levels attained by several proposed procedures, for known $\nu = \sigma/\mu$ (but unknown μ and σ), and normally distributed demand. Among these proposals is the *exact rule,* that leads to the required service levels exactly. In Section 5, ν is unknown too, which leads to time-varying safety factors; these *advanced rules* are based on the results of Section 3.

In Section 6 the demand distribution is changed to gamma. Since all previously suggested procedures prove to be insufficient, an *adaptive rule* incorporating a feedback mechanism is proposed. Section 7 concludes the paper with conclusions, discussions and ideas for further research.

2. Fully known normal demand distribution

Let t denote a review moment, when the inventory position is measured and - if necessary - replenished to a certain level S_t. Let demand X_t during the subsequent review period satisfy

$$X_t \sim N(\mu, \sigma^2)$$

where, for the moment, μ and σ^2 are assumed to be given. Stock-out occurs when X_t exceeds S_t; hence the P_1-service criterion can be written as

$$P(X_t \leq S_t) = P_1 \tag{2.1}$$

It is easily seen that this equality is satisfied by the choice of a constant order-up-to-level level, S_1 say, given by

$$S_1 = \mu + \sigma \Phi^{-1}(P_1) \tag{2.2}$$

where Φ is the standardnormal distribution function. (Note that a slight complication is introduced by the normality assumption: for large σ in particular, X_t can take negative values, implying that S_1 may be *less* than $S_1 - X_t$, the value of the net stock at the end of the corresponding review period. We solve this mainly theoretical problem by assuming that the surplus stock is sent back to the supplier.)

If $Y^+ = \max(Y, 0)$, the P_2-service criterion can be written as

$$\frac{E\left[(X_t - S_t)^+\right]}{\mu} = 1 - P_2$$

Again, for S_t a constant, S_2 say, can be taken. Let ϕ denote the density of a standardnormal variable Z and define the function G on $\mathbb{R}$ by

$$G(k) = E\left[(Z-k)^{+}\right] = \int_{k}^{\infty} (z-k)\,\phi(z)\,dz$$

Then the P_2-service criterion may be rewritten as

(2.2) $$\frac{\sigma}{\mu} G\left(\frac{S_2-\mu}{\sigma}\right) = 1-P_2$$

Now, the well-known properties

$$\phi'(k) = -k\,\phi(k), \qquad G(k) = \phi(k) - k\,\Phi(-k)$$

imply

$$G'(k) = -\Phi(-k) < 0$$

Hence, G is decreasing on $\mathbb{R}$ and G^{-1} exists, so that an explicit expression for S_2 can be obtained:

$$S_2 = \mu + \sigma G^{-1}\left[(1-P_2)/\nu\right]$$

where $\nu = \sigma/\mu$ is the coefficient of variation. All these results are standard; e.g., see Silver et al. (1998).

Introducing the notation

(2.3) $$\begin{cases} c_1 = \Phi^{-1}(P_1) \\ c_2 = G^{-1}\left[(1-P_2)/\nu\right] \end{cases}$$

the order-up-to-levels for the two service criteria can be written as one formula:

(2.4) $$S_i = \mu + c_i\sigma, \quad i = 1, 2$$

where S_i relates to the P_i-criterion. Table 2.1 presents some values of c_1 (depending on P_1) and c_2 (depending on P_2 and ν). Note that c_2 is increasing in ν. For the tabulated values of P_i, the case $c_1 = c_2$ occurs for $\nu = 1.10, 2.23, 2.40,$ and 2.65, respectively.

Table 2.1 Values of c_i.

c_2 ν	P_i	0.90	0.925	0.95	0.975
0.5		0.493	0.671	0.902	1.256
0.75		0.741	0.902	1.115	1.443
1		0.902	1.055	1.256	1.569
1.25		1.021	1.167	1.360	1.663
1.5		1.115	1.256	1.443	1.738
c_1		1.282	1.440	1.645	1.960

3. Stationary normal demand distribution with unknown parameter(s)

Of much more practical importance is the case of unknown μ and σ. Let M_t and D_t be estimators of μ and σ, respectively, based on observations up to time t. Then, the most straightforward idea is to plug these estimators into (2.4); for the P_i-service criterion this leads to the (time-varying) order-up-to-level

$$S_t = M_t + c_i D_t$$

This method is advocated throughout the standard literature on inventory control, thereby at least suggesting that the required service level P_i is (approximately) attained. However, Silver and Rahnama (1986, 1987) showed that for a (s, Q)-control policy with a cost criterion the cost penalties are not symmetric around the true reorder point; consequently, they explored the possibility of deliberately biasing upwards the safety factor. In a previous discussion paper (Strijbosch et al., 1997), we explained for the (R, S)-policy that the service level, attained by this standard procedure, falls short of P_1 uniformly - even in the most simple setting. We expand this reasoning here.

Assume only μ unknown and take $\sigma = 1$ for simplicity. In this stationary situation, the best estimator for μ at moment t is the sample mean

$$\overline{X} = \frac{1}{t}\sum_{i=0}^{t-1} X_i$$

Hence, according to the literature, the order-up-to-level at time t should equal

$$S_t = \overline{X} + c_i$$

Then, $\overline{X} \sim N(\mu, 1/t)$ implies for the P_1-criterion

$$P(X_t > S_t) = P(X_t - \overline{X} > c_1) = 1 - \Phi(c_1/\sqrt{1+1/t}) \;\; > 1 - \Phi(c_1) = 1 - P_1$$

Although $E(S_t) = \mu + c_1 = S_1$, the attained service level falls short with respect to the desired level P_1. This effect is due to the asymmetry of the distribution of X_t around $\mu + c_1$; by consequence, positive and negative deviations of S_t around its mean have a different impact: $P(X_t > \mu + c_1 - \varepsilon) + P(X_t > \mu + c_1 + \varepsilon) > 2(1 - P_1)$ for all $\varepsilon > 0$. See Figure 1.

As to the P_2-criterion, it follows in this setting

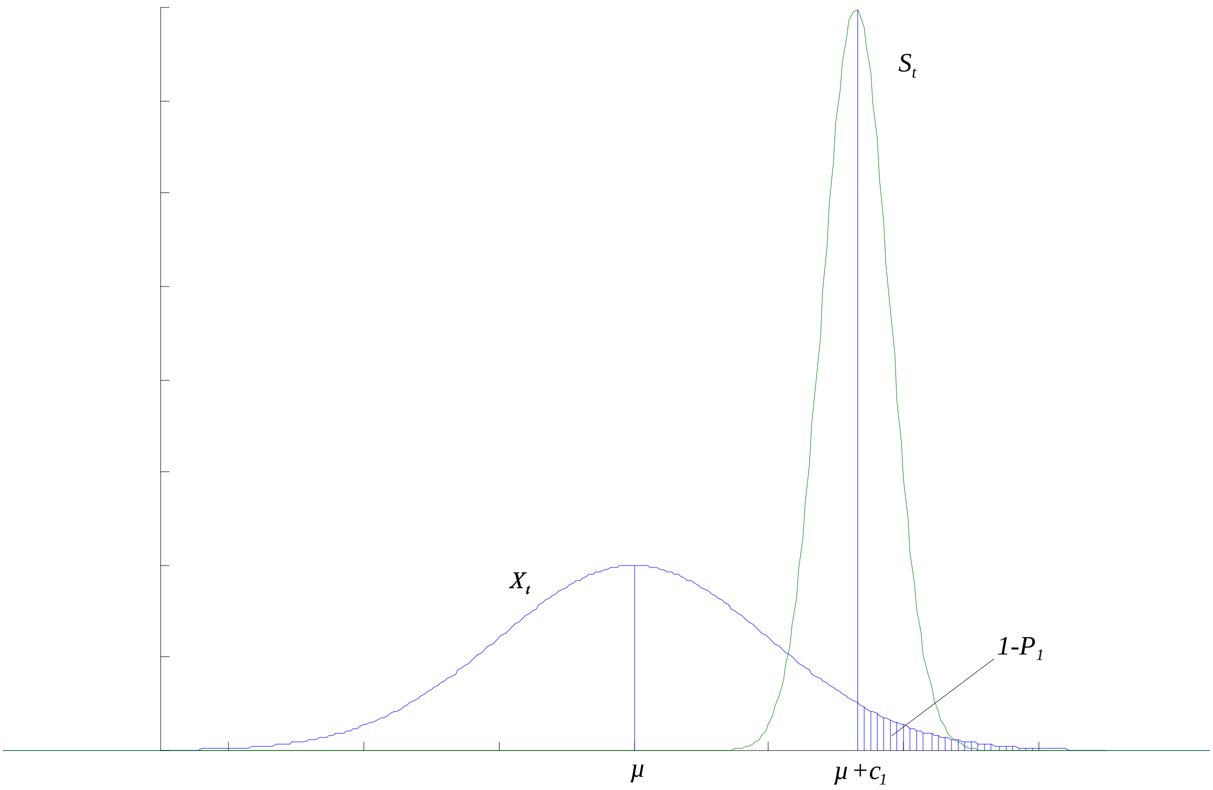
$$E\left[(X_t - S_t)^+\right] = \sqrt{1+1/t}\; G\left(c_2/\sqrt{1+1/t}\right)$$

Since G is decreasing, this implies

$$\frac{E\left[(X_t - S_t)^+\right]}{\mu} > \frac{G(c_2)}{\mu} = 1 - P_2$$

So indeed, the service levels attained by the generally advocated replenishment quantity $S_t = M_t + c_i D_t$ are too low - even in this most ideal situation.

Figure 1. The distribution of X_t and S_t.



The conclusion is that the safety factors c_i should be increased; hence order-up-to-levels of the form

(3.1) $\quad S_t = M_t + k\,D_t$

will be considered in the sequel. Here, k will depend on the service criterion chosen, and on the estimators for μ and σ. As the notation indicates, constants k are searched for; however, in Sections 5 and 6, k will be allowed to be time-dependent.

In finding the value of k, two approaches will be followed. First, we will derive *approximate* values k_i, corresponding with the two P_i-service criteria, by assuming a normal distribution for S_t. Since in fact the distribution of S_t is rather complicated, values r_i of k leading to *exact* service levels P_i will be calculated as well.

If μ is estimated by a - weighted or unweighted - mean of (past) observations, M_t is normally distributed. Further, the usual sample variance has a χ^2-distribution, which for large sample size can be approximated by a normal distribution. Consequently, the distribution of S_t will not be too far from normal in this case. A similar argument holds for other estimators of σ, e.g. if simple exponential smoothing (SES) is used. Hence, to allow an approximate analytical derivation of the constant safety factors k in (3.1), we *assume* for the moment that S_t has a normal distribution - although this is not true in fact.

So, we assume that $S_t \sim N(\mu_s, \sigma_s^2)$ where

$$(3.2) \quad \begin{cases} \mu_s = \mu + kE(D_t) = \mu + k\mu_d \\ \sigma_s^2 = V(M_t) + 2k\,Cov(M_t, D_t) + k^2 V(D_t) \end{cases}$$

Write for simplicity $W_t = X_t - S_t$, so that $W_t \sim N(-k\mu_d, \sigma_w^2)$ with $\sigma_w^2 = \sigma^2 + \sigma_s^2$. Using k_i now, it follows:

$$P\left(W_t \le 0\right) = \Phi\left(k_1 \mu_d / \sigma_w\right)$$

$$E\left(W_t^+\right) = \sigma_w G\left(k_2 \mu_d / \sigma_w\right)$$

An immediate consequence is that the constants k_i, corresponding with the two P_i-service criteria, in this case should satisfy

$$(3.3) \quad \begin{cases} k_1 = \Phi^{-1}(P_1)\,\sigma_w / \mu_d \\ k_2 = G^{-1}\left[(1-P_2)\,\mu / \sigma_w\right] \sigma_w / \mu_d \end{cases}$$

These two expressions may be compared with (2.3). Since $\sigma < \sigma_w$,

$$G^{-1}\left[(1-P_2)\,\mu / \sigma_w\right] > G^{-1}\left[(1-P_2)\,\mu / \sigma\right] = c_2$$

holds. For an unbiased estimator D_t^2 of σ^2, Jensen's inequality gives

$$\mu_d = E(D_t) < \sqrt{E(D_t^2)} = \sigma < \sigma_w$$

Consequently,

$$k_i > c_i, \; i = 1,2$$

holds: in case of unknown μ and σ, the safety factors c_i should be increased indeed.

Note that this can be achieved implicitly by using a positively biased estimator D_t for σ; this agrees with the advise in Jacobs and Wagner (1989) to multiply MAD by a larger factor than 1.25 when estimating σ.

From the results of Silver and Rahnama (1986, 1987), who investigated an (s,Q)-policy with a cost criterion, and of Strijbosch et al. (1997), who studied a (R,S)-policy with a P_1-criterion only, Silver et al. (1998) concluded , that "it may be advisable to somewhat scale up $\hat{\sigma}_L$ (here: D_t) to reflect the added variability due to the statistical estimation of σ_L (here: σ)." This conclusion appears to hold as well when applying the P_2-criterion. The numerical consequences of nevertheless using the safety factors c_i when μ and σ have to be estimated, are presented in the next two sections.

Note that the k_i can not be calculated directly from (3.3). The main complication is that $\sigma_w^2 = \sigma_s^2 + \sigma^2$ itself depends on k_i, according to (3.2). In a Monte Carlo environment using only stationary demand data and a well-defined procedure to determine M_t and D_t, it is possible to find k_i iteratively by substituting an initial guess for k_i, calculating sample estimates for μ, σ_w and μ_d from the entire simulation run, applying (3.3) for the next guess of k_i, and so on until convergence occurs. These k_i would lead exactly to the prescribed service levels, *if* S_t were normally distributed; therefore, in a sense, the service levels attained using k_i show the deviations of normality of S_t.

The exact distribution of S_t can be approximated numerically for any given value of

ν. From this exact distribution, the values of k in (3.1) exactly corresponding with the prescribed P_i-value can be determined by a simple search procedure; these *exact* safety factors will be denoted by r_i. The same simulation runs as before can be used.

The next section will clarify how the Monte Carlo study has been conducted and will present the exact rule r_i, as well as the approximating values k_i.

4. Simulation study

The goal of the Monte Carlo experiment is to establish numerically that application of the standard rule in general yields a performance level that is lower than the desired level, as was proved in Section 3. Furthermore, we want to find the exact rule r_i and to find, and investigate the behaviour of the approximating rule k_i.

Basicly, a time series $\{X_t\}$ is generated where t runs from -1,000 to 10,000, using the part -1,000 to 0 as an initialisation of the corresponding procedures. For all simulations we assume a normally distributed demand with $\mu = 10$; five different values of the coefficient of variation ν $(0.5,\ 0.75,\ 1.0,\ 1.25,\ 1.5)$ are used. In the simulation program, we obtained M_t and D_t by applying SES to the five different samples:

$$\begin{cases} M_t = \alpha X_t + (1-\alpha) M_{t-1} \\ MAD_t = \omega \left| X_t - M_{t-1} \right| + (1-\omega) MAD_{t-1} \\ D_t = 1.25\, MAD_t \end{cases}$$

Although it is possible to obtain a smoothed variance by using $V_t = \omega (X_t - M_{t-1})^2 + (1-\omega) V_{t-1}$, we choose to link up with the common use of the mean absolute deviation. Simulation results (not presented here) indicate that the main conclusions in this paper would not be affected by the use of the smoothed variance. Further, note that D_t, at least for normally distributed demands, is an unbiased estimator of the standard deviation of the forecast error, and a biased estimator of σ: appendix A implies $V(X_t - M_{t-1}) = \frac{\alpha}{2-\alpha}\sigma^2 \neq \sigma^2 = V(X_t)$. However, substituting σ by an estimator of the forecast error standard deviation makes sense, since we have to account for the variability of X_t as well as for the variability of the forecast itself, as advocated throughout the standard literature.

On each of the five samples we applied SES with all combinations of $\alpha \in \{0.01,\ 0.05,\ 0.10,\ 0.15\}$ and $\omega \in \{0.01,\ 0.03,\ 0.06,\ 0.09\}$. Simulation results are obtained with all (5*4*4=) 80 series for each $P_i \in \{0.90,\ 0.925,\ 0.95,\ 0.975\}$. From these complete series, the safety factors r_i leading to the exact service levels P_i were calculated. Similarly, the approximate safety factors k_i, based on normality of S_t, were found iteratively from (3.3).

Table 4.1. *Deviations* $1000(r_1/c_1 - 1)$ *(and* $1000(r_1/k_1 - 1)$, *within parentheses); normally distributed demand.*

$1000 * P_1$		900	925	950	975
α	ω				
0.01	0.01	-2 (-10)	6 (-2)	12 (2)	30 (18)
	0.03	10 (-3)	13 (-2)	11 (-8)	38 (14)
	0.06	20 (-1)	23 (-1)	30 (-1)	47 (5)
	0.09	30 (2)	32 (-3)	49 (5)	68 (6)
0.05	0.01	-8 (-13)	4 (-3)	16 (7)	20 (9)
	0.03	5 (-6)	8 (-6)	29 (11)	30 (6)
	0.06	9 (-10)	24 (1)	35 (4)	47 (5)
	0.09	19 (-8)	39 (5)	44 (1)	65 (3)
0.10	0.01	-9 (-14)	-1 (-7)	16 (8)	21 (12)
	0.03	-5 (-15)	9 (-3)	27 (10)	27 (4)
	0.06	6 (-11)	24 (1)	36 (7)	47 (7)
	0.09	16 (-10)	36 (3)	47 (5)	62 (1)
0.15	0.01	-12 (-16)	1 (-4)	13 (7)	16 (7)
	0.03	1 (-9)	6 (-6)	16 (0)	23 (2)
	0.06	5 (-12)	18 (-4)	25 (-4)	43 (2)
	0.09	15 (-11)	25 (-7)	43 (0)	67 (7)

For brevity, only the deviations of c_i and k_i from the exact rule r_i are shown; with the aid of Table 2.1, the exact value r_i can be recalculated. Consider the P_1-criterion first. The service level $P_1 = 0.975$ is reached for $c_1 = 1.960$, if μ and σ are known. For unknown parameters, this level was reached exactly for $r_1 = 2.091$ in a certain case. Hence, for unknown parameters, c_1 should be increased by 67‰. This figure can be found in the lower righthand corner of Table 4.1. Note that 68 is the highest number in this table, indicating that the standard safety factor c_1 falls short at most 7%. In the same case, the value $k_1 = 2.076$ was found, 7‰ below the exact value $r_1 = 2.091$. This permillage can be found between parentheses in Table 4.1. In summary: Table 4.1 shows to what extent the safety factors c_1 should be increased, in case of

Table 4.2. *Deviations from prescribed service level* $1000\left[P_1 - \hat{P}(c_1)\right]$ *(and* $1000\left[P_1 - \hat{P}(k_1)\right]$ *within parentheses); normally distributed demand.*

$1000 * P_1$		900	925	950	975
α	ω				
0.01	0.01	-1 (-2)	1 (0)	2 (0)	4 (2)
	0.03	2 (0)	3 (-1)	2 (-1)	5 (1)
	0.06	5 (0)	4 (0)	4 (0)	6 (0)
	0.09	6 (0)	6 (-1)	6 (0)	8 (1)
0.05	0.01	-1 (-2)	1 (0)	3 (1)	2 (1)
	0.03	1 (-1)	2 (-2)	4 (2)	4 (1)
	0.06	3 (-2)	5 (0)	6 (1)	5 (1)
	0.09	4 (-1)	7 (1)	8 (0)	7 (0)
0.10	0.01	-2 (-3)	0 (-2)	3 (1)	2 (1)
	0.03	-1 (-4)	2 (0)	4 (1)	3 (1)
	0.06	1 (-3)	5 (0)	6 (2)	5 (1)
	0.09	2 (-3)	6 (1)	8 (1)	6 (0)
0.15	0.01	-3 (-4)	0 (-1)	2 (1)	3 (1)
	0.03	0 (-2)	1 (-1)	2 (0)	3 (0)
	0.06	1 (-3)	4 (-1)	3 (0)	6 (0)
	0.09	3 (-2)	5 (-1)	6 (0)	7 (1)

unknown μ and σ. The exact values r_1 should be used for stationary normal demand. Note that the necessary correction increases with ω and P_1. The approximate value k_1 is nearly as good as r_1; for $P_1 = 0.9$ or 0.925, k_1 may be too high. Finally, note that the results in Table 4.1 are indifferent for the coefficient of variation ν. This is a consequence of the fact that c_1 is independent of ν, as the distribution of all relevant statistics (like S_t and W_t) in this case belong to a location-scale family; cf. Strijbosch et al. (1997).

Table 4.2 shows the consequences of using c_1 or k_1 in stead of r_1: it presents the deviations of the desired service level P_1 (again in permillages) . The actually attained service levels are denoted by $\hat{P}(c_1)$ and $\hat{P}(k_1)$, respectively. The number 7 in the lower righthand corner means that c_1 leads to service level 0.968 in stead of 0.975. This looks not too serious, but note that the stock-out probability is increased with 28%: from 2.5 to 3.2%. Note that there is a clear effect of ω; as ω increases, the performance of the standard rule tends to grow worse, independent of the value of α. In Appendix A an expression is established for $V(V_t)$, which turns out to depend mainly on ω; this could be a partial explanation of the minor dependence of the results in Tables 4.1 and 4.2 on α. The second conclusion from Table 4.2 is that $\hat{P}(k_1)$ approaches the desired service level quite good; there is a tendency of overcompensating for smaller service levels, in agreement with our conclusions from Table 4.1.

The analysis for the P_2-criterion is similar, although the outcomes now depend on ν (assumed to be known). From Table 4.3 the safety factors r_2 can be found, leading exactly to the desired service level P_2. The necessary increase of c_2 is of the same magnitude as in the P_1-case. Again, k_2 is a good approximation of r_2. Table 4.4 shows to what extent the actually attained service levels $\hat{P}(c_2)$ and $\hat{P}(k_2)$ fall short. As in Table 4.2, use of the standard safety factor c_2 may lead to an increase of the prescribed stock-out probability of 28%. The dependence on ν appears to be only minor.

Table 4.3. *Deviations* $1000(r_2/c_2-1)$ *(and* $1000(r_2/k_2-1)$*, within parentheses); normally distributed demand.*

ν		0.5			1			1.5		
$1000{*}P_2$		900	950	975	900	950	975	900	950	975
α	ω									
0.01	0.01	0 (-8)	11 (1)	14 (3)	9 (-1)	13 (2)	18 (6)	9 (-1)	15 (4)	21 (7)
	0.03	2 (-10)	16 (1)	20 (2)	14 (0)	19 (2)	26 (4)	15 (0)	24 (3)	31 (6)
	0.06	6 (-8)	21 (1)	30 (3)	20 (0)	29 (2)	41 (4)	25 (2)	36 (3)	47 (5)
	0.09	8 (-8)	28 (1)	41 (4)	27 (0)	41 (3)	57 (4)	34 (2)	50 (3)	65 (3)
0.05	0.01	12 (-8)	12 (-1)	16 (4)	10 (-3)	15 (3)	20 (7)	11 (-1)	18 (5)	20 (6)
	0.03	14 (-8)	17 (-2)	24 (4)	16 (-3)	23 (3)	28 (4)	18 (-1)	26 (4)	29 (2)
	0.06	18 (-6)	23 (-2)	34 (4)	22 (-3)	33 (3)	42 (3)	27 (-1)	39 (3)	46 (2)
	0.09	22 (-6)	30 (-2)	45 (4)	29 (-3)	45 (3)	57 (1)	36 (-1)	52 (3)	64 (0)
0.10	0.01	30 (-6)	17 (-3)	18 (2)	16 (-4)	18 (2)	22 (6)	13 (-4)	21 (5)	20 (5)
	0.03	32 (-4)	22 (-3)	25 (2)	21 (-4)	25 (1)	30 (4)	21 (-3)	29 (4)	30 (2)
	0.06	37 (-6)	28 (-4)	36 (2)	27 (-5)	35 (1)	44 (2)	29 (-3)	41 (3)	46 (0)
	0.09	39 (-6)	34 (-4)	47 (2)	33 (-5)	46 (1)	58 (1)	38 (-3)	54 (2)	64 (-2)
0.15	0.01	47 (-6)	22 (-5)	21 (0)	21 (-6)	20 (-1)	24 (5)	18 (-4)	24 (4)	22 (4)
	0.03	49 (-6)	28 (-5)	29 (0)	27 (-6)	28 (-1)	32 (2)	25 (-4)	30 (1)	33 (2)
	0.06	53 (-6)	33 (-6)	39 (0)	32 (-7)	38 (-1)	46 (1)	34 (-3)	42 (0)	51 (1)
	0.09	57 (-4)	40 (-6)	49 (-1)	39 (-7)	49 (-2)	62 (1)	42 (-5)	55 (-1)	69 (-1)

Summarizing this section: we found the exact safety factors r_i for stationary normal demand with known $\nu = \sigma/\mu$; besides, approximations k_i to r_i were presented. All in all, the consequences of using in stead of these the standard values c_i of (2.3) are not too dramatic.

Table 4.4: *Deviations from prescribed service level* $1000\left[P_2 - \hat{P}(c_2)\right]$ *(and* $1000\left[P_2 - \hat{P}(k_2)\right]$ *, within parentheses); normally distributed demand.*

ν		0.5			1			1.5		
$1000{*}P_2$		900	950	975	900	950	975	900	950	975
α	ω									
0.01	0.01	0 (-1)	1 (0)	1 (0)	1 (0)	2 (0)	2 (1)	2 (0)	2 (1)	2 (1)
	0.03	0 (-1)	1 (0)	1 (0)	2 (0)	3 (0)	2 (0)	4 (0)	4 (1)	3 (0)
	0.06	0 (-1)	2 (0)	2 (0)	3 (0)	4 (0)	4 (0)	5 (0)	6 (1)	5 (1)
	0.09	1 (-1)	2 (0)	3 (0)	4 (0)	5 (0)	5 (0)	7 (0)	8 (1)	7 (0)
0.05	0.01	1 (-1)	1 (0)	1 (0)	2 (-1)	2 (0)	2 (1)	2 (0)	3 (1)	2 (1)
	0.03	1 (-1)	1 (0)	2 (0)	3 (-1)	3 (0)	3 (0)	4 (0)	4 (1)	3 (0)
	0.06	1 (-1)	2 (0)	2 (0)	4 (0)	4 (0)	4 (0)	6 (0)	6 (1)	5 (0)
	0.09	2 (0)	2 (0)	3 (0)	5 (0)	6 (0)	5 (0)	8 (0)	8 (0)	7 (0)
0.10	0.01	2 (-1)	1 (0)	1 (0)	2 (-1)	2 (0)	2 (1)	3 (-1)	3 (1)	2 (1)
	0.03	2 (-1)	2 (0)	2 (0)	3 (-1)	3 (0)	3 (0)	5 (-1)	5 (1)	3 (0)
	0.06	3 (0)	2 (0)	2 (0)	4 (-1)	5 (0)	4 (0)	6 (-1)	6 (0)	5 (0)
	0.09	3 (0)	3 (0)	3 (0)	5 (-1)	6 (0)	5 (0)	8 (-1)	8 (0)	7 (0)
0.15	0.01	4 (-1)	2 (0)	1 (0)	4 (-1)	3 (0)	2 (1)	4 (-1)	4 (1)	2 (0)
	0.03	4 (-1)	2 (0)	2 (0)	4 (-1)	4 (0)	3 (0)	6 (-1)	5 (0)	4 (0)
	0.06	4 (0)	3 (0)	3 (0)	5 (-1)	5 (0)	4 (0)	7 (-1)	7 (0)	5 (0)
	0.09	4 (0)	3 (0)	3 (0)	6 (-1)	6 (0)	6 (0)	9 (-1)	9 (0)	7 (0)

5. Time-varying safety factors; normally distributed demand

Up to now, constant safety factors were looked for, corresponding to a stationary situation. For given value of k, the distribution of S_t in (3.1) was simulated; from these distributions the necessary value r_i of k could be found for given ν. Similarly, the value of σ_w and μ_d in (3.3) were calculated from an entire simulation run; the approximate safety factors k_i followed, assuming normality of S_t.

In practice of course, no run of 10,000 stationary observations X_t is available; besides,

ν will be unknown: therefore, the value of k in (3.3) has to be determined from t previous observations in a varying environment. A straightforward generalization is to infer information on σ_w and μ_d with SES as well, and substitute the corresponding estimates sw_t and md_t in (3.3). So we obtain time-varying safety factors $k_{i,t}$:

$$(5.1) \qquad \begin{cases} k_{1,t} = \Phi^{-1}(P_1)\, sw_t / md_t \\ k_{2,t} = G^{-1}\left[(1-P_2)\, M_t / sw_t\right] sw_t / md_t \end{cases}$$

We will refer to (5.1) as the *advanced rule*. The plugged-in estimates are determined according to the following exponential smoothing rules:

$$\begin{cases} mw_t = \alpha_w W_t + (1-\alpha_w)\, mw_{t-1} \\ MADw_t = \omega_w \left| W_t - mw_{t-1} \right| + (1-\omega_w)\, MADw_{t-1} \\ sw_t = 1.25\, MADw_t \\ md_t = \alpha_d D_t + (1-\alpha_d)\, md_{t-1} \end{cases}$$

In order to keep the number of combinations of input parameters limited, we used constant values of the additional smoothing parameters: $\alpha_w = \omega_w = \alpha_d = 0.01$. In the same spirit, the *standard rule* (2.3) should be modified: since ν is unknown, c_2 must become time-varying:

$$(5.2) \qquad \begin{cases} c_1 = \Phi^{-1}(P_1) \\ c_{2,t} = G^{-1}\left[(1-P_2)\, M_t / D_t\right] \end{cases}$$

Table 5.1 shows the performance of using safety factor $c_{2,t}$ versus $k_{2,t}$ for various values of P_2, ν, α and ω. The main observations from this table are as follows. First of all, small values of α and ω yield performance levels close to the desired level, as should be the case. Using the standard rule does not lead to dramatic deviations from P_2 (at most 1% with a few exceptions), and when ω increases a decrease of the realised service level is observed.

Whenever the standard rule yields a performance under P_2, the advanced rule produces a performance closer to the desired one. Whenever the standard rule yields a performance higher than P_2, the advanced rule too produces a performance which is almost always higher than P_2, but less higher.

Table 5.2 shows to what extent the *mean* safety factor $c_{2,t}$ must increase if μ, σ and ν are unknown. The table shows, that, especially for small values of α, $c_{2,t}$ should be corrected upwards (up to 7%) in order to obtain the desired service level.

Table 5.1: *Deviations of desired service level* $1000\left[P_2 - \hat{P}(c_{2,t})\right]$ *(and* $1000\left[P_2 - \hat{P}(k_{2,t})\right]$, *within parentheses); normally distributed demand.*

ν		0.5			1			1.5		
$1000{*}P_2$		900	950	975	900	950	975	900	950	975
α	ω									
0.01	0.01	0 (0)	1 (0)	1 (1)	2 (1)	2 (1)	2 (1)	2 (1)	3 (2)	2 (2)
	0.03	1 (-1)	2 (0)	2 (0)	3 (1)	3 (1)	3 (1)	5 (1)	4 (2)	4 (2)
	0.06	2 (-1)	3 (0)	3 (0)	5 (1)	5 (1)	5 (1)	8 (1)	7 (1)	6 (1)
	0.09	2 (-1)	4 (0)	4 (0)	7 (0)	8 (1)	7 (1)	11 (1)	11(1)	9 (1)
0.05	0.01	-1 (-1)	0 (0)	1 (0)	-2 (0)	0 (1)	1 (1)	-3 (-1)	-1 (1)	1 (1)
	0.03	0 (-1)	1 (0)	1 (0)	0 (0)	1 (0)	2 (1)	-1 (-1)	1 (1)	2 (1)
	0.06	0 (-1)	2 (0)	2 (0)	2 (-1)	4 (0)	4 (1)	3 (-1)	5 (0)	4 (1)
	0.09	1 (-1)	3 (0)	4 (0)	4 (-1)	6 (0)	6 (0)	6 (-1)	8 (0)	7 (0)
0.10	0.01	-3 (-1)	-1 (0)	0 (0)	-6 (-2)	-3 (0)	-1 (1)	-10 (-4)	-4 (-1)	-2 (0)
	0.03	-2 (-1)	0 (0)	1 (0)	-4 (-2)	-1 (0)	0 (1)	-8 (-4)	-3 (-1)	0 (0)
	0.06	-1 (-1)	1 (-1)	2 (0)	-2 (-2)	1 (0)	2 (0)	-4 (-5)	1 (-1)	2 (0)
	0.09	0 (-1)	2 (-1)	3 (0)	0 (-2)	3 (-1)	4 (0)	0 (-5)	4 (-2)	5 (-1)
0.15	0.01	-4 (-2)	-2 (-1)	1 (0)	-10 (-3)	-5 (-1)	-2 (0)	-17 (-9)	-8 (-3)	-4 (-1)
	0.03	-4 (-2)	-2 (-1)	0 (0)	-9 (-3)	-4 (-1)	-1 (0)	-15 (-9)	-7 (-3)	-3 (-1)
	0.06	-3 (-2)	0 (-1)	1 (0)	-6 (-3)	-1 (-1)	1 (0)	-11 (-10)	-3 (-4)	0 (-1)
	0.09	-2 (-2)	1 (-1)	2 (0)	-4 (-4)	1 (-1)	3 (0)	-7 (-10)	0 (-4)	2 (- 2)

We did not tabulate the corresponding results for the P_1-case as they are not very different from those in Tables 4.1 and 4.2. The general conclusions are that the advanced rule yields a better performance than the standard rule, which means a closer approximation to the desired service level. However, the gain is not impressive.

Table 5.2: *Deviations* $1000(r_2/\bar{c}_{2,t}-1)$ *(and* $1000(r_2/\bar{k}_{2,t}-1)$*, within parentheses); normally distributed demand.*

ν		0.5			1			1.5		
1000* P_2		900	950	975	900	950	975	900	950	975
α	ω									
0.01	0.01	6 (-2)	13 (4)	16 (6)	12 (6)	15 (6)	19 (8)	12 (5)	17 (8)	22 (11)
	0.03	14(-4)	20 (3)	24 (4)	20 (5)	23 (5)	29 (6)	20 (4)	26 (7)	33 (8)
	0.06	27 (-4)	29 (2)	36 (2)	29 (3)	35 (3)	45 (4)	33 (4)	40 (4)	51 (4)
	0.09	35 (-4)	39 (1)	50 (1)	39 (2)	49 (2)	62 (1)	43 (3)	56 (2)	70 (-1)
0.05	0.01	2 (4)	7 (4)	13 (8)	1 (8)	10 (11)	16 (13)	-1 (8)	10 (14)	13 (13)
	0.03	8 (2)	13 (3)	22 (7)	9 (8)	19 (11)	25 (10)	8 (8)	19 (12)	23 (8)
	0.06	20 (2)	23 (2)	35 (5)	19 (5)	32 (9)	40 (7)	20 (6)	33 (8)	41 (5)
	0.09	33 (2)	33 (1)	48 (3)	29 (4)	45 (7)	57 (3)	31 (3)	48 (5)	61 (1)
0.10	0.01	-2 (18)	1 (9)	9 (12)	-9 (16)	2 (18)	10 (20)	-19 (7)	-3 (17)	2 (17)
	0.03	6 (16)	9 (9)	17 (10)	-1 (15)	11 (17)	20 (17)	-10 (6)	7 (14)	12 (12)
	0.06	16 (16)	18 (7)	30 (9)	8 (13)	23 (14)	35 (14)	0 (2)	20 (9)	29 (7)
	0.09	26 (16)	28 (5)	43 (7)	17 (11)	36 (13)	50 (9)	10 (-2)	34 (5)	48 (1)
0.15	0.01	-6 (30)	-3 (15)	5 (16)	-19 (22)	-6 (25)	4 (27)	-41 (-4)	-18 (14)	-9 (18)
	0.03	0 (28)	4 (15)	14 (15)	-12 (22)	3 (24)	13 (23)	-32 (-6)	-11 (9)	2 (14)
	0.06	12 (30)	13 (13)	26 (13)	-3 (20)	15 (21)	28 (19)	-21 (-11)	3 (3)	20 (8)
	0.09	24 (30)	23 (12)	39 (11)	6 (17)	27 (19)	46 (17)	-11 (-15)	17 (-3)	39 (2)

6. Gamma demand distribution

Although the normal distribution is a very common model for demand, it is instructive to look at another demand distribution as well. One of the most usual and attractive aternatives is the gamma distribution: like demand, it takes only positive values and is skewed to the right. Further, note that it is the base under the Mixed-Erlang distribution, advocated by Tijms (1994) a.o. Hence, in this section, we start by applying our previously described methods to the case of gamma distributed demand.

Consider again the general expression (3.1) for order-up-to-levels: $S_t = M_t + k\,D_t$. As before, for different values of k, the distribution of S_t can be simulated, now for gamma distributed demand. From these simulated distributions, the values of k can be calculated for given ν, leading to the exact service levels P_i; these exact rules are denoted by r_i again. Presenting them is postponed until Table 6.6.

Table 6.1: *Deviations of desired service level applying the advanced rule; gamma distributed demand.*

	$100\left[P_1 - \hat{P}(k_{1,t})\right]$				$100\left[P_2 - \hat{P}(k_{2,t})\right]$			
$100 * P_i$	90	92.5	95	97.5	90	92.5	95	97.5
ν								
0.5	1	1	2	2	1	1	2	2
0.75	1	1	2	2	3	4	4	3
1	1	2	2	3	6	7	7	6
1.25	1	2	3	3	10	10	10	8
1.5	1	2	3	3	13	13	12	10

Several approximations to r_i were discussed in the previous sections. Using the safety factors c_i in (2.3) now has obvious disadvantages, since they were explicitly based upon normality. The same holds for the k_i in (3.3) and the time-varying counterparts $c_{2,t}$, $k_{i,t}$.

Nevertheless, studying their performance in the present situation is very instructive: it shows the consequences of assuming normality, when in fact demand is gamma distributed. Only a few results are shown here; they regard the safety factors $k_{i,t}$ that performed best in the normal case. Table 6.1 shows the extent to which the realised service level deviates from the desired level. The figures are averages over the values of α and ω used throughout. Note that the deviations are presented as percentages in this section, rather than as permillages like before! Comparison with the numbers between parentheses in Tables 4.2 and 5.1 shows that the obtained service levels are much lower now - not surprisingly, of course. E.g. for $\nu = 1.5$ and $P_2 = 90\%$, the actually attained service level was very close to 90% for normally distributed demand, while it is only 77% in case of the gamma distribution.

Therefore, we looked for still another approximating rule, allowing more flexibility in the assumptions. In fact, we suggest to use empirically obtained information on the attained service level in the recent past to adaptively tuning the safety factors; that is, we try to learn from the past as a compensation for the deficiencies of the applied formulae. To be more precise, we apply the advanced rule but take as safetyfactors $q_{i,t} = q_i k_{i,t}$ where the advanced rule $k_{i,t}$ serves as initial guess and the multiplier q_i is updated after each n periods. Updating is done according to the attained service level over the past n periods (thus, individual updates of q_i are based on non-overlapping intervals), as follows:

$$(6.1)\qquad q_i = \begin{cases} q_i * (1+\delta_u) & \text{for} \quad P_i - \hat{P}_i(q_{i,t}) > \Delta \\ q_i & \text{for} \quad |P_i - \hat{P}_i(q_{i,t})| < \Delta \\ q_i / (1+\delta_d) & \text{for} \quad P_i - \hat{P}_i(q_{i,t}) < -\Delta \end{cases}$$

where t is some integer multiple of n and $\hat{P}_i(q_{i,t})$ is measured over the last n periods before period t. We call (6.1) the *adaptive rule*; it is studied by means of simulation.

Again, in order to keep the number of combinations of input parameters limited, we took only a few values of the extra parameters. Firstly, we set $\Delta = 0.01$ throughout. As we realise that the parameter n for the adaptive rule is crucial in the performance of this rule, we use two levels: $n = 60$ for circumstances where the level of stationarity and information allows it, and $n = 20$ where this is not, or much less, the case. As the demand distribution is skewed to the right, it makes sense to correct the safety factor asymmetrically. This leads us to investigate the combination $\delta_u = \delta_d = 0.05$ (for $n = 20$), as well as $\delta_u = 0.075$ and $\delta_d = 0.05$ (for $n = 60$). In the simulation program a precautionary upper bound for S_t had to be set, since there are circumstances where $q_i k_{i,t}$ tends to grow to fast yielding excessive order-up-to levels. An upperbound of 100 was chosen (recall that $\mu = 10$); in practice this level should be set by management, e.g. indicated by physical inventory restrictions.

Tables 6.2 and 6.3 present the obtained service levels for $n = 20$. The outcomes turned out to be not very sensitive to the values of α and ω; hence, averaging over these two smoothing constants has been applied again. For comparison, the corresponding service levels for the standard rule are given between brackets.

Table 6.2: *Deviations of desired service level* $100\left[P_1 - \hat{P}(q_{1,t})\right]$ *(and* $100\left[P_1 - \hat{P}(c_1)\right]$ *within parentheses); gamma distributed demand;* $\delta_u = \delta_d = 0.05$, $\Delta = 0.01$, $n = 20$.

$100 * P_1$ / ν	90	92.5	95	97.5
0.5	0 (1)	0 (2)	0 (2)	1 (2)
0.75	1 (2)	1 (2)	0 (3)	1 (3)
1	0 (2)	0 (3)	0 (3)	1 (4)
1.25	1 (2)	1 (3)	0 (4)	1 (4)
1.5	0 (2)	1 (3)	1 (4)	1 (5)

Table 6.3: *Deviations of desired service level* $100\left[P_2 - \hat{P}(q_{2,t})\right]$ *(and* $100\left[P_2 - \hat{P}(c_{2,t})\right]$ *within parentheses); gamma distributed demand;* $\delta_u = \delta_d = 0.05$, $\Delta = 0.01$, $n = 20$.

$100 * P_2$ / ν	90	92.5	95	97.5
0.5	0 (1)	0 (2)	1 (2)	1 (2)
0.75	1 (4)	1 (4)	1 (4)	2 (4)
1	2 (7)	2 (7)	2 (7)	3 (7)
1.25	3 (11)	3 (11)	3 (11)	4 (10)
1.5	4 (14)	4 (14)	5 (14)	5 (13)

Our adaptive rule falls short at most 1% for the P_1-criterion, and at most 5% for P_2. It is a clear improvement of the standard rule, and even of the advanced rule (compare Table 6.1).

Table 6.4: *Deviations* $100(r_1/\bar{q}_{1,t}-1)$ *(and* $100(r_1/c_1-1)$ *within parentheses); gamma distributed demand;* $\delta_u=\delta_d=0.05$, $\Delta=0.01$, $n=20$.

ν \ $100*P_1$	90	92.5	95	97.5
0.5	0 (9)	1 (14)	1 (19)	8 (27)
0.75	2 (10)	5 (18)	2 (27)	6 (39)
1	1 (13)	1 (22)	1 (37)	8 (59)
1.25	3 (16)	2 (29)	2 (45)	8 (70)
1.5	3 (19)	6 (32)	1 (48)	10 (82)

Table 6.5: *Deviations* $100(r_2/\bar{q}_{2,t}-1)$ *(and* $100(r_2/\bar{c}_{2,t}-1)$ *within parentheses); gamma distributed demand;* $\delta_u=\delta_d=0.05$, $\Delta=0.01$, $n=20$.

ν \ $100*P_2$	90	92.5	95	97.5
0.5	7 (23)	7 (26)	8 (30)	11 (36)
0.75	11 (41)	11 (46)	13 (52)	22 (63)
1	13 (70)	16 (77)	20 (87)	35 (103)
1.25	18 (97)	23 (105)	31 (117)	50 (136)
1.5	25 (123)	31 (135)	41 (150)	65 (175)

Tables 6.4 and 6.5 show to what extent the adaptive rule falls short of the exact rule. So, for $P_2=0.95$ and $\nu=1$, the adaptive rule yields a safety factor that must be multiplied with 1.20 to reach the desired service level. It is clear that the standard rule produces in general far too low safety factors. Although this adaptive rule is a major improvement, its corrective strength, especially for the P_2 criterion, is not large enough. If $n=60$ is used and *asymmetrically* adjusting terms $\delta_u=0.075$ and $\delta_d=0.05$, the adaptive rule yields deviations of the desired

service level of at most 1%, even for the P_2 criterion. This better performance is mainly due to the larger value of n and less to the asymmetric correction. Of course, such a large value of n would only be feasible in practice when demand data is reasonably stationary over a (very) long time interval. However, the results with $n=20$ and $n=60$ indicate that routinely using recent performance information could be worthwhile.

Table 6.6 summarizes the rules discussed up to now. It contains the averaged values (over α and ω) of the mean safety factors c_1, $c_{2,t}$, $k_{i,t}$, $q_{i,t}(20)$, $q_{i,t}(60)$ and r_i, where $q_{i,t}(20)$ and $q_{i,t}(60)$ are obtained with $n=20$, $\delta_u=\delta_d=0.05$ and $n=60$, $\delta_u=0.075$, $\delta_d=0.05$, respectively. This table clearly demonstrates the effect of using a more elaborated procedure for establishing the safety factor, and of the value of information used.

Table 6.6: *Safety factors for the two service criteria; gamma distributed demand.*

P_i	ν	c_1	$\bar{k}_{1,t}$	$\bar{q}_{1,t}(20)$	$\bar{q}_{1,t}(60)$	r_1	$\bar{c}_{2,t}$	$\bar{k}_{2,t}$	$\bar{q}_{2,t}(20)$	$\bar{q}_{2,t}(60)$	r_2
0.9	0.5	1.282	1.329	1.390	1.401	1.395	0.477	0.486	0.547	0.559	0.587
	0.75	1.282	1.341	1.382	1.411	1.410	0.715	0.732	0.909	0.984	1.010
	1	1.282	1.355	1.434	1.447	1.446	0.867	0.893	1.297	1.418	1.471
	1.25	1.282	1.380	1.444	1.491	1.491	0.972	1.017	1.622	1.887	1.913
	1.5	1.282	1.415	1.478	1.544	1.523	1.042	1.109	1.863	2.211	2.324
0.925	0.5	1.440	1.502	1.627	1.643	1.637	0.656	0.672	0.772	0.797	0.826
	0.75	1.440	1.519	1.617	1.679	1.693	0.879	0.906	1.155	1.246	1.280
	1	1.440	1.540	1.738	1.762	1.761	1.021	1.062	1.567	1.729	1.811
	1.25	1.440	1.575	1.820	1.899	1.862	1.120	1.187	1.876	2.259	2.300
	1.5	1.440	1.628	1.791	1.948	1.898	1.187	1.287	2.136	2.580	2.788
0.95	0.5	1.645	1.733	1.925	1.976	1.951	0.889	0.916	1.070	1.127	1.155
	0.75	1.645	1.757	2.043	2.071	2.086	1.093	1.137	1.470	1.610	1.665
	1	1.645	1.792	2.227	2.298	2.259	1.225	1.292	1.904	2.179	2.291
	1.25	1.645	1.843	2.324	2.415	2.377	1.317	1.422	2.178	2.741	2.857
	1.5	1.645	1.932	2.401	2.518	2.436	1.379	1.542	2.440	3.121	3.450
0.975	0.5	1.960	2.101	2.309	2.526	2.492	1.244	1.299	1.522	1.681	1.696
	0.75	1.960	2.140	2.562	2.805	2.719	1.424	1.508	1.910	2.245	2.328
	1	1.960	2.211	2.882	3.190	3.111	1.541	1.673	2.324	2.942	3.135
	1.25	1.960	2.301	3.092	3.428	3.340	1.624	1.823	2.563	3.468	3.836
	1.5	1.960	2.468	3.248	3.663	3.558	1.680	2.000	2.797	3.945	4.626

7. Conclusions and further research

In practical inventory control, demand distribution parameters have to be estimated. Since demand is usually non-stationary, exponential smoothing procedures are mostly applied. For a (R,S)-control policy with zero lead times, we showed both analytically and by simulation that combining these estimates with the standard safety factors c_i does not lead to the required service levels, even when demand is stationary and normal. This conclusion holds for the P_1- as well as the P_2-service criterion. An additional theoretical argument for this phenomenon can be derived from Appendix B. As in Section 3, we find for $W_t = X_t - S_t$ the properties

$$(7.1) \qquad \begin{cases} P\left(W_t \le 0\right) = \Phi\left(\dfrac{c_1 + d}{g}\right) \\ E\left(W_t^+\right) = g\,\sigma\, G\left(\dfrac{c_2 + d}{g}\right) \end{cases}$$

where $\mu_s = \mu + (c_i + d)\,\sigma$ and $\sigma_s^2 = (g^2 - 1)\,\sigma^2$, $(g > 1)$. Formula (7.1) implies that even when $d = 0$ (the order-up-to-level is unbiasedly estimated), the attained service will be too low, and that positively biased order-up-to-levels may compensate for the effect resulting from $g > 1$. These conclusions formally hold for normal demand only.

Extensive Monte Carlo investigations reveal the quantitative effects under several circumstances that are relevant for practitioners. In all cases a constant safety factor r_i is determined which would lead exactly to the desired service level. This exact rule r_i is then used as a yardstick to measure the extent to which the safety factors obtained via several other procedures should be increased in order to better approach the desired service level. The main results of these investigations are the following.

For normally distributed demand, the standard rule yields service levels that fall short not more that 0.7%, generally. However, comparing the deviations with the *failure levels* $1 - P_i$ gives a quite different picture of these outcomes: the failure level may be increased by

28%! Hence, we advocate an increase of the standard safety factors c_i with 5%, in contrast to leading textbooks like Silver et al. (1998) and Fogarty et al. (1991).

Our advanced rule offers a further refinement: by using additional SES-estimates, a time-varying approximation to the exact rule is obtained that proved to be quite adequate for normally distributed demand. Of course, for known ν, the exact safety factor r_i themselves should be used - both in case of normally and gamma distributed demand.

In the latter case, our approximations are not good enough. Therefore, an adaptive rule has been constructed that employs a feedback mechanism based on the attained service level in the recent past. This rule is able to produce service levels which come much closer to the desired ones. Probably, however not tested, this rule may compensate for other anomalous circumstances, like some forms of non-stationarity, as well. Note that our adaptive rule is a more general concept than the rolling-horizon concept (cf. Silver et al. (1998)). In fact, we added in this paper to the rolling horizon approach a feedback mechanism which accounts for deviations of the assumptions.

Although our appraoch was general in the sense that two service criteria were considered, as well as two usual demand distributions, it still has serious limitations. In our view, the most important draw-back is the zero lead-time assumption. Dropping this assumption leads to dependence between demand during consecutive review plus lead-time periods, thereby complicating the analysis. Nevertheless, we will concentrate on this issue in the near future.

Another straightforward generalisation is to study other inventory control systems than (R,S). Further, modelling demand by means of the normal distribution - although quite usual - has serious draw-backs, as indicated before. Other demand distributions are much more appropriate, like the gamma distributions used here as well. It is to be expected that, as the demand distribution, chosen for modelling, better fits observations, the performance of the combined forecasting-inventory control approach will fall short less. Limited research (results not presented) applying the mixed-Erlang distribution (cf. Tijms (1994)) however, indicates that an unacceptable gap between attained and desired service levels remains.

And, finally, our proposals were tested only in stationary situations. As these are virtually non-existent in practice, extensive investigations in non-stationary environments are needed.

Acknowledgement

We are indebted to Ruud Heuts for his careful reading of the manuscript.

References

Brown, R.G. (1963)

Smoothing, forecasting and prediction of discrete time series,

Prentice-Hall, Englewood Cliffs, N.J.

Eppen, G.D., and Martin, R.K. (1988)

Determining safety stock in the presence of stochastic lead time and demand,

Management Science 34, 1380-1390.

Fogarty, D.W., Blackstone, J.H., and Hoffmann, T.R. (1991)

Production and inventory management, South-Western.

Gardner, E.S. (1990)

Evaluating forecast performance in an inventory control system,

Management Science 36, 490-499.

Jacobs, R.A., and Wagner H.M. (1989)

Reducing inventory system costs by using robust demand estimators,

Management Science 35, 771-787.

Karmarkar, U.S. (1994)

A robust forecasting technique for inventory and leadtime management,

Journal of Operations Management 12, 45-54.

Silver, E.A., and Smith, D.J. (1981)

Setting individual item production rates under significant lead time conditions,

INFOR 15, 1-19.

Silver, E.A., Pyke, D.F., and Peterson, R. (1998)

Inventory management and production planning and scheduling,

Wiley, New York.

Silver E.A., and Rahnama, M.R. (1986)

The cost effects of statistical sampling in selection the reorder point in a common

inventory model,
Journal of the Operational Research Society 37, 705-713.

Silver E.A., and Rahnama, M.R. (1987)
Biased selection of the inventory reorder point when demand parameters are statistically estimated,
Engineering Costs and Production Economics 12, 283-292.

Strijbosch, L.W.G., Heuts, R.M.J. and van der Schoot, E.H.M. (1998)
Improved spare parts inventory management: a case study,
Submitted to the Journal of the Operational Research Society ,
Center discussion paper 98135, Tilburg University, The Netherlands.

Strijbosch, L.W.G., Moors, J.J.A., de Kok, A.G. (1997)
On the interaction between forecasting and inventory control,
Reserach Memorandum FEW742, Tilburg University, The Netherlands.

Tijms, H.C. (1994)
Stochastic models: an algorithmic approach,
Wiley, Chichester.

Watson, R.B. (1987)
The effects of demand-forecast fluctuations on customer service and inventory cost when demand is lumpy, Journal of the Operational Research Society 38, 75-82.

APPENDIX A: Derivation of $V(V_t)$.

The forecast update formula for the variance of the forecast error is given by

$$V_t = \omega E_t^2 + (1-\omega) V_{t-1} = \omega \sum_{i=0}^{\infty} (1-\omega)^i E_{t-i}^2$$

where

$$E_t = X_t - M_{t-1}$$

Denote $\sigma_E^2 = E(V_t)$. Because of stationarity, V_t and V_{t-1} are identically distributed. Using the update formula and $\sigma_V^2 = V(V_t) = V(V_{t-1})$ we may write:

$$\sigma_V^2 = 2\omega^2\sigma_E^4 + (1-\omega)^2\sigma_V^2 + 2\omega(1-\omega)C(E_t^2, V_{t-1}),$$

or

$$(2-\omega)\sigma_V^2 = 2\omega\sigma_E^4 + 2\omega\sum_{j=1}^{\infty} (1-\omega)^j C(E_t^2, E_{t-j}^2). \qquad \text{(a)}$$

Using the well-known result

$$E(Y_1Y_2Y_3Y_4) = \sigma_{12}\sigma_{34} + \sigma_{13}\sigma_{24} + \sigma_{14}\sigma_{23}$$

for a multivariate normal distributed variable $Y = (Y_1, Y_2, Y_3, Y_4)$ with $E(Y) = 0$ and variance-covariance matrix $\Sigma = \{\sigma_{ij}\}_{i=1,..,4;\, j=1,..4}$, we may establish

$$C(E_t^2, E_{t-j}^2) = 2\,C^2(E_t, E_{t-j}). \qquad \text{(b)}$$

An expression for $C(E_t, E_{t-j})$ can be found in a few consecutive steps.

$$C(M_{t-i}, X_{t-j}) = \begin{cases} \alpha(1-\alpha)^{j-i}\sigma^2 & , \ j \ge i \\ 0 & , \ j < i \end{cases}$$

$$C(M_{t-i}, M_{t-j}) = \frac{\alpha}{2-\alpha}(1-\alpha)^{|i-j|}\sigma^2$$

$$C(M_{t-i}, E_{t-j}) = \begin{cases} \dfrac{\alpha}{2-\alpha}(1-\alpha)^{j-i}\sigma^2 & , \ j \ge i \\ \dfrac{-\alpha}{2-\alpha}(1-\alpha)^{i-j-1}\sigma^2 & , \ j < i \end{cases}$$

$$C(E_{t-i},X_{t-j})=\begin{cases} 0 \ , \ j<i \\ \sigma^2 \ , \ j=i \\ -\alpha(1-\alpha)^{j-i-1}\sigma^2 \ , \ j>i \end{cases}$$

$$C(E_{t-i},E_{t-j})=\begin{cases} \dfrac{-\alpha}{2-\alpha}(1-\alpha)^{|i-j|-1}\sigma^2 \ , \ j\neq i \\ \dfrac{2}{2-\alpha}\sigma^2=\sigma_E^2 \ , \ j=i \end{cases}$$

Thus,

$$C(E_t,E_{t-j})=\frac{-\alpha}{2-\alpha}(1-\alpha)^{j-1}\sigma^2, \text{ for } j\geq 1$$

and with (a) and (b) this gives

$$\sigma_V^2=\frac{\omega\sigma_E^4}{2-\omega}\left\{2+\frac{\alpha^2(1-\omega)}{1-(1-\omega)(1-\alpha)^2}\right\}.$$

Note that $\dfrac{\alpha^2(1-\omega)}{1-(1-\omega)(1-\alpha)^2}\leq\dfrac{1}{7}$ for $\alpha\leq 0.25$ so that σ_V^2 may be approximated by $\dfrac{2\omega\sigma_E^4}{2-\omega}$ in most practical cases.

Assuming $\mu=0$ for simplicity, another result can easily be obtained:

$$C(X_t,\ X_sX_u)=0\ (\forall_{s,t,u})\ \Rightarrow\ C(M_t,\ E_{t-i}^2)=0\ (\forall_{i\geq 0})\ \Rightarrow\ C(M_t,V_t)=0.$$

In other words, there is no correlation between the forecast and the variance of the forecast error.

Note that the covariance formulae in this appendix include variance formulae (for $i=j$); some of these can be found elsewhere, e.g. in Brown (1963).

APPENDIX B

Let U and V be independent and standard normal distributed variables ($U \sim N(0,1)$ and $V \sim N(0,1)$). Then $W = aU + b \;\sim\; N(b, a^2)$, and $Z = V - W \;\sim\; N(-b, 1 + a^2)$.

Now the following expressions can be established:

B.1
$$\int_{-\infty}^{\infty} \phi(u)\left(1 - \Phi(au+b)\right)du = \int_{-\infty}^{\infty} \phi(u) P\left(V \geq aU + b \mid U = u\right) du = P\left(V \geq aU + b\right) =$$
$$= P(Z \geq 0) = P\left(\frac{Z+b}{\sqrt{1+a^2}} \geq \frac{b}{\sqrt{1+a^2}}\right) = 1 - \Phi\left(\frac{b}{\sqrt{1+a^2}}\right)$$

B.2
$$\int_{-\infty}^{\infty} \phi(u)\, G(au+b)\, du = \int_{-\infty}^{\infty} \phi(u) E\left[\left(V - (aU+b) \mid U = u\right)^{+}\right] du =$$
$$= E\left[\left(V - (aU+b)\right)^{+}\right] = E\left[Z^{+}\right] = \sqrt{1+a^2}\, G\left(\frac{b}{\sqrt{1+a^2}}\right)$$

Silver and Smith (1981) already elaborated expression B.2 in a more circumstantial way. Silver et al. (1998) mention both expressions, however the first expression contains an error.